Differential analysis of matrix convex functions

Frank Hansen and Jun Tomiyama

January 12, 2006

Revised June 12, 2006

Abstract

We analyze matrix convex functions of a fixed order defined in a real interval by differential methods as opposed to the characterization in terms of divided differences given by Kraus [7]. We obtain for each order conditions for matrix convexity which are necessary and locally sufficient, and they allow us to prove the existence of gaps between classes of matrix convex functions of successive orders, and to give explicit examples of the type of functions contained in each of these gaps. The given conditions are shown to be also globally sufficient for matrix convexity of order two. We finally introduce a fractional transformation which connects the set of matrix monotone functions of each order n with the set of matrix convex functions of the following order $n+1$.

1 Introduction

Let f be a real function defined in an interval I. It is said to be n-convex if

$$f(\lambda A+(1-\lambda)B)\le \lambda f(A)+(1-\lambda)f(B)\qquad \lambda\in[0,1]$$

for arbitrary Hermitian $n\times n$ matrices A and B with spectra in I. It is said to be n-concave if $-f$ is n-convex, and it is said to be n-monotone if

$$A\le B\quad\Rightarrow\quad f(A)\le f(B)$$

for arbitrary Hermitian $n\times n$ matrices A and B with spectra in I. We consider the interval of definition to be part of the specification of a function. The notions of n-convexity and n-monotonicity are therefore not associated solely

with an assignment rule, but may depend on the interval in which the rule is applied. We denote by $P_n(I)$ the set of n-monotone functions defined in an interval I, and by $K_n(I)$ the set of n-convex functions defined in I.

We shall sometimes use the standard regularization procedure, cf. for example Donoghue [3, Page 11]. Let φ be a positive and even C^∞-function defined on the real axis, vanishing outside the closed interval $[-1,1]$ and normalized such that

$$\int_{-1}^{1} \varphi(x)\,dx = 1.$$

For any locally integrable function f defined in an open interval (a,b) we form its regularization

$$f_\epsilon(t) = \frac{1}{\epsilon}\int_a^b \varphi\left(\frac{t-s}{\epsilon}\right) f(s)\,ds \qquad t \in \mathbf{R}$$

for small $\epsilon > 0$, and realize that it is infinitely many times differentiable. For $t \in (a+\epsilon, b-\epsilon)$ we may also write

$$f_\epsilon(t) = \int_{-1}^{1} \varphi(s) f(t-\epsilon s)\,ds.$$

If f is continuous, then f_ϵ converges uniformly to f on any compact subinterval of (a,b). If in addition f is n-convex (or n-monotone) in (a,b), then f_ϵ is n-convex (or n-monotone) in the slightly smaller interval $(a+\epsilon, b-\epsilon)$. Since the pointwise limit of a sequence of n-convex (or n-monotone) functions is again n-convex (or n-monotone), we may therefore in many applications assume that an n-convex or n-monotone function is sufficiently many times differentiable.

We remind the reader [3, Chapter VII Theorem VI] that a 2-monotone function defined in an open interval automatically is continuously differentiable. The first statement of the following lemma can also be found in the same reference.

Lemma 1.1. *Let I be an open interval and consider a real function f defined in I.*

(1) If f is 2-monotone and $f'(t_0) = 0$ for a single $t_0 \in I$, then f is a constant and therefore $f'(t) = 0$ for all $t \in I$.

(2) If f is twice continuously differentiable, 2-convex and $f''(t_0) = 0$ for a single $t_0 \in I$, then f is affine and therefore $f''(t) = 0$ for all $t \in I$.

Proof. We use the characterizations of 2-monotonicity by Löwner [9] and 2-convexity by Kraus [7]. If f is 2-monotone then

$$\det\begin{pmatrix}[t_0,t_0]_f & [t_0,t]_f\\ [t,t_0]_f & [t,t]_f\end{pmatrix} = f'(t_0)f'(t) - [t_0,t]_f^2 \geq 0$$

from which the first statement follows. Similarly, if f is 2-convex then

$$\det\begin{pmatrix}[t_0,t_0,t_0]_f & [t_0,t_0,t]_f\\ [t_0,t,t_0]_f & [t_0,t,t]_f\end{pmatrix} = \frac{1}{2}f''(t_0)[t_0,t,t]_f - [t_0,t_0,t]_f^2 \geq 0$$

from which the second statement follows.

QED

1.1 Differential conditions

If I is open and f is twice differentiable we introduce for real numbers $s, t_1, \dots, t_n$ in I the leading determinants

$$D_r(s) = \det\begin{pmatrix}[t_1,s,t_1]_f & [t_1,s,t_2]_f & \cdots & [t_1,s,t_r]_f\\ [t_2,s,t_1]_f & [t_2,s,t_2]_f & \cdots & [t_2,s,t_r]_f\\ \vdots & \vdots & & \vdots\\ [t_r,s,t_1]_f & [t_r,s,t_2]_f & \cdots & [t_r,s,t_r]_f\end{pmatrix} \tag{1}$$

for $r = 1, \dots, n$. Kraus proved [7] that f is n-convex if and only if the matrix

$$H(s) = \Big([t_i,s,t_j]_f\Big)_{i,j=1}^n \tag{2}$$

is positive semi-definite for every sequence $t_1, \dots, t_n \in I$ and $s = t_1, \dots, t_n$.

Theorem 1.2. *Let f be a real $2n$ times continuously differentiable function defined in an open interval I. The determinants defined in (1) can be written*

$$D_r(s) = \det M_r \cdot \prod_{k=1}^{r-1}\prod_{l=1}^{r-k}(t_{k+l} - t_l)^2 \qquad r = 1, \dots, n$$

where $M_r = \big(m_{ij}\big)_{i,j=1}^r$ and $m_{ij} = [t_1, \dots, t_i, s, t_1, \dots, t_j]_f$. If in addition f is n-convex, then the matrix

$$K_n(f;t) = \left(\frac{f^{i+j}(t)}{(i+j)!}\right)_{i,j=1}^n$$

is positive semi-definite for each $t \in I$. On the other hand, if $K_n(f;t_0)$ is positive definite for some $t_0 \in I$, then f is n-convex in some open interval J with $t_0 \in J \subseteq I$.

Proof. Let $t_1, \ldots, t_n$ be n distinct points in the interval I. Since there is no possibility of confusion, we shall omit the reference to the function f in the divided differences. For each $r = 2, \ldots, n$ we intend to prove

$$D_r(s) = \det M_r(p) \cdot \prod_{k=1}^{p} \prod_{l=1}^{r-k} (t_{k+l} - t_l)^2 \tag{3}$$

by induction for $p = 1, \ldots, r-1$ where

$$M_r(p) = \begin{pmatrix} M_p & \big([t_1, \ldots, t_i, s, t_{j-p}, \ldots, t_j]\big)_{i,j} \\ \big([t_{i-p}, \ldots, t_i, s, t_1, \ldots, t_j]\big)_{i,j} & \big([t_{i-p}, \ldots, t_i, s, t_{j-p}, \ldots, t_j]\big)_{i,j} \end{pmatrix}.$$

Note that $M_r(p)$ is an $r \times r$ matrix written as a block matrix with the $p \times p$ matrix M_p as the $(1,1)$ block entry. The indices i and j refer to the absolute row and column numbers in $M_r(p)$. The row index i in block entry $(2,1)$ hence runs from $p+1$ to r, and the column index j runs from 1 to p. When (3) is proved, the first part of the theorem follows by setting $p = r-1$ and noting that $M_r(r-1) = M_r$.

In the determinant expression (1) we subtract the first row from the second row, the second row from the third and so forth until the $(r-1)$th row is subtracted from the rth row. We thus obtain

$$D_r(s) =$$

$$\begin{vmatrix} [t_1, s, t_1] & [t_1, s, t_2] & \cdots & [t_1, s, t_r] \\ [t_2, s, t_1] - [t_1, s, t_1] & [t_2, s, t_2] - [t_1, s, t_2] & \cdots & [t_2, s, t_r] - [t_1, s, t_r] \\ \vdots & \vdots & & \vdots \\ [t_r, s, t_1] - [t_{r-1}, s, t_1] & [t_r, s, t_2] - [t_{r-1}, s, t_2] & \cdots & [t_r, s, t_r] - [t_{r-1}, s, t_r] \end{vmatrix}$$

and since for $i = 2, \ldots, r$ and $j = 1, \ldots, r$ the difference

$$\begin{aligned} [t_i, s, t_j] - [t_{i-1}, s, t_j] &= [t_i, s, t_j] - [s, t_j, t_{i-1}] = (t_i - t_{i-1})[t_i, s, t_j, t_{i-1}] \\ &= (t_i - t_{i-1})[t_{i-1}, t_i, s, t_j], \end{aligned}$$

we obtain the expression

$$D_r(s) = (t_2 - t_1)(t_3 - t_2) \cdots (t_r - t_{r-1}) \times \begin{vmatrix} [t_1, s, t_1] & [t_1, s, t_2] & \cdots & [t_1, s, t_r] \\ [t_1, t_2, s, t_1] & [t_1, t_2, s, t_2] & \cdots & [t_1, t_2, s, t_r] \\ \vdots & \vdots & & \vdots \\ [t_{r-1}, t_r, s, t_1] & [t_{r-1}, t_r, s, t_2] & \cdots & [t_{r-1}, t_r, s, t_r] \end{vmatrix}.$$

We then subtract the first column from the second column, the second column from the third and so forth until the $(r-1)$th column is subtracted from the rth column and obtain

$$D_r(s) = (t_2-t_1)^2(t_3-t_2)^2\cdots(t_r-t_{r-1})^2\times \begin{vmatrix} [t_1,s,t_1] & [t_1,s,t_1,t_2] & \cdots & [t_1,s,t_{r-1},t_r] \\ [t_1,t_2,s,t_1] & [t_1,t_2,s,t_1,t_2] & \cdots & [t_1,t_2,s,t_{r-1},t_r] \\ \vdots & \vdots & & \vdots \\ [t_{r-1},t_r,s,t_1] & [t_{r-1},t_r,s,t_1,t_2] & \cdots & [t_{r-1},t_r,s,t_{r-1},t_r] \end{vmatrix}$$

since for $i=1$ and $j=2,\dots,r$ the difference

$$\begin{aligned}[t_1,s,t_j]-[t_1,s,t_{j-1}] &= [t_j,t_1,s]-[t_1,s,t_{j-1}] = (t_j-t_{j-1})[t_j,t_1,s,t_{j-1}] \\ &= (t_j-t_{j-1})[t_1,s,t_{j-1},t_j],\end{aligned}$$

while for $i=2,\dots,r$ and $j=2,\dots,r$ the difference

$$\begin{aligned}&[t_{i-1},t_i,s,t_j]-[t_{i-1},t_i,s,t_{j-1}] = [t_j,t_{i-1},t_i,s]-[t_{i-1},t_i,s,t_{j-1}] \\ &= (t_j-t_{j-1})[t_j,t_{i-1},t_i,s,t_{j-1}] = (t_j-t_{j-1})[t_{i-1},t_i,s,t_{j-1},t_j].\end{aligned}$$

Note that the above expression proves (3) for $p=1$ and any $r=2,\dots,n$. In particular (3) is valid for $r=2$.

Assume now that $r\geq 3$ and (3) is valid for some $p\leq r-2$. We subtract, in the matrix $M_r(p)$, the $(p+1)$th row from the $(p+2)$th row, the $(p+2)$th row from the $(p+3)$th row until the $(r-1)$th row is subtracted from the rth row and obtain

$$\det M_r(p) = (t_{p+2}-t_1)(t_{p+3}-t_2)\cdots(t_r-t_{r-(p+1)})\times$$
$$\begin{vmatrix} M_p & \big([t_1,\dots,t_i,s,t_{j-p},\dots,t_j]\big)_{i,j} \\ \big([t_1,\dots,t_{p+1},s,t_1,\dots,t_j]\big)_j & \big([t_1,\dots,t_{p+1},s,t_{j-p},\dots,t_j]\big)_j \\ \big([t_{i-(p+1)},\dots,t_i,s,t_1,\dots,t_j]\big)_{i,j} & \big([t_{i-(p+1)},\dots,t_i,s,t_{j-p},\dots,t_j]\big)_{i,j} \end{vmatrix}.$$

This is so since for $i=p+2,\dots,r$ and $j=1,\dots,p$ the difference

$$\begin{aligned}&[t_{i-p},\dots,t_i,s,t_1,\dots,t_j]-[t_{i-(p+1)},\dots,t_{i-1},s,t_1,\dots,t_j] \\ &= (t_i-t_{i-(p+1)})[t_{i-(p+1)},t_{i-p},\dots,t_i,s,t_1,\dots,t_j]\end{aligned}$$

and similarly

$$\begin{aligned}&[t_{i-p},\dots,t_i,s,t_{j-p},\dots,t_j]-[t_{i-(p+1)},\dots,t_{i-1},s,t_{j-p},\dots,t_j] \\ &= (t_i-t_{i-(p+1)})[t_{i-(p+1)},t_{i-p},\dots,t_i,s,t_{j-p},\dots,t_j].\end{aligned}$$

Note that row $p+1$ was left unchanged.

We finally subtract, in the above determinant expression, the $(p+1)$th column from the $(p+2)$th column, the $(p+2)$th column from the $(p+3)$th column until the $(r-1)$th column is subtracted from the rth column and obtain

$$\det M_r(p) = (t_{p+2}-t_1)^2(t_{p+3}-t_2)^2 \cdots (t_r - t_{r-(p+1)})^2 \times$$
$$\begin{vmatrix} M_{p+1} & \big([t_1,\ldots,t_i,s,t_{j-(p+1)},\ldots,t_j]\big)_{i,j} \\ \big([t_{i-(p+1)},\ldots,t_i,s,t_1,\ldots,t_j]\big)_{i,j} & \big([t_{i-(p+1)},\ldots,t_i,s,t_{j-(p+1)},\ldots,t_j]\big)_{i,j} \end{vmatrix}$$

by calculations as above. Our calculations show that

$$\det M_r(p) = (t_{p+2}-t_1)^2(t_{p+3}-t_2)^2 \cdots (t_r - t_{r-(p+1)})^2 \det M_r(p+1)$$

and consequently

$$D_r(s) = \det M_r(p+1) \cdot \prod_{k=1}^{p+1} \prod_{l=1}^{r-k} (t_{k+l}-t_l)^2$$

which shows (3) by induction.

The second statement of the theorem now follows by choosing $s=t$ and letting all the numbers $t_1,\ldots,t_n$ tend to t.

If f is n-convex, then it follows by Kraus' theorem that the matrix $H(s)$ is positive semi-definite, thus all the leading determinants (1) are non-negative. Since the numbers $t_1,\ldots,t_n$ are distinct, it follows that also the determinants of the matrices M_r are non-negative for $r=1,\ldots,n$. By choosing $s=t$ and letting all the numbers $t_1,\ldots,t_n$ tend to t, we derive that the leading principal determinants of the matrix $K_n(f;t)$ are all non-negative. But since each principal submatrix of $K_n(f;t)$ in this way may be obtained as a leading principal submatrix by first making a suitable joint permutation of the rows and columns in $H(s)$, it follows that the determinants of all principal submatrices of $K_n(f;t)$ are non-negative. Therefore $K_n(f;t)$ is indeed positive semi-definite.

Finally, if the matrix $K_n(f;t_0)$ is positive definite in some point $t_0 \in I$, we use that the entries $[s,t_i,t_j]_f$ of the matrix $H(s)$ are continuous functions of t_i,t_j and s to obtain that the matrix $H(s)$ is positive definite for all $t_1,\ldots,t_n$ and $s=t_1,\ldots,t_n$ in an open interval J with $t_0 \in J \subseteq I$. The assertion now follows from the characterization by Kraus [7] of matrix convexity in terms of divided differences. **QED**

1.2 The existence of gaps

Proposition 1.3. *Let I be a finite interval, and let m and n be natural numbers with $m \geq 2n$. There exists an n-concave and n-monotone polynomial $f_m\colon I \to \mathbf{R}$ of degree m. Likewise there exists an n-convex and n-monotone polynomial $g_m\colon I \to \mathbf{R}$ of degree m.*

Proof. We may without loss of generality assume that I is an open interval and then obtain the statement of the proposition for other finite interval types by considering restrictions of a polynomial defined on an open interval. The interval I may thus be written on the form $I = (t_0 - c, t_0 + c)$ for some $t_0 \in \mathbf{R}$ and a positive real number c. We introduce the polynomial p_m of degree m given by

$$p_m(t) = b_1 t + b_2 t^2 + \cdots + b_m t^m$$

where

$$b_k = \int_{-1}^{0} t^{k-1}\,dt = \frac{(-1)^{k-1}}{k}$$

for $k = 1, \ldots, m$. Thus the pth derivative $p_m^{(p)}(0) = p! \cdot b_p$ for $p = 1, \ldots, 2n$ and consequently

$$M_n(p_m;0) = \left(\frac{p_m^{(i+j-1)}(0)}{(i+j-1)!}\right)_{i,j=1}^{n} = (b_{i+j-1})_{i,j=1}^{n},$$

where we used the notation in [3, 5]. Similarly

$$K_n(p_m;0) = \left(\frac{p_m^{(i+j)}(0)}{(i+j)!}\right)_{i,j=1}^{n} = (b_{i+j})_{i,j=1}^{n}.$$

Take a vector $c = (c_1, \ldots, c_n) \in \mathbf{C}^n$, then

$$(M_n(p_m;0)c \mid c) = \sum_{i,j=1}^{n} b_{i+j-1} c_j \bar{c}_i = \int_{-1}^{0} \left| \sum_{i=1}^{n} c_i t^{i-1} \right|^2 dt$$

and

$$(K_n(p_m;0)c \mid c) = \sum_{i,j=1}^{n} b_{i+j} c_j \bar{c}_i = \int_{-1}^{0} t \left| \sum_{i=1}^{n} c_i t^{i-1} \right|^2 dt.$$

Since the coefficients in a polynomial are all zero if the polynomial is the zero function, we derive that $M_n(p_m;0)$ is positive definite and $K_n(p_m;0)$ is

negative definite. Hence there exists an $\alpha > 0$ such that $M_n(p_m; t)$ is positive definite and $K_n(p_m; t)$ is negative definite in the interval $(-\alpha, \alpha)$, thus p_m is n-monotone and n-concave on $(-\alpha, \alpha)$. The polynomial

$$f_m(t) = p_m(\alpha c^{-1}(t - t_0)) \qquad t \in I$$

then has the desired properties. The second statement is proved by choosing the coefficients

$$b_k = \int_0^1 t^{k-1}\, dt = \frac{1}{k}$$

and then follow the same steps as in the above proof. **QED**

Proposition 1.4. *Let I be an interval, and let $n \geq 2$ be a natural number. There are no n-convex polynomials of degree m in I for $m = 3, \ldots, 2n - 1$.*

Proof. If f_m is an n-convex polynomial of degree m in I and t_0 is an inner point in I, then

$$p_m(t) = f_m(t - t_0)$$

is n-convex in a neighborhood of zero and may be written on the form

$$p_m(t) = b_0 + b_1 t + \cdots + b_m t^m$$

where $b_m \neq 0$. We calculate the derivatives

$$p_m^{(m-1)}(0) = (m-1)!\, b_{m-1}, \quad p_m^{(m)}(0) = m!\, b_m, \quad p_m^{(m+1)}(0) = 0.$$

If m is even and thus of the form $m = 2l$ for some $l \geq 2$, then the principal submatrix of $K_n(p_m; 0)$ consisting of the rows and columns with numbers $l - 1$ and $l + 1$ is given by

$$\begin{pmatrix} b_{m-2} & b_m \\ b_m & 0 \end{pmatrix},$$

and it has determinant $-b_m^2 < 0$. If m is odd and thus of the form $m = 2l + 1$ for some $l \geq 1$, then the principal submatrix of $K_n(p_m; 0)$ consisting of the rows and columns with numbers l and $l + 1$ is given by

$$\begin{pmatrix} b_{m-1} & b_m \\ b_m & 0 \end{pmatrix},$$

and this matrix also has determinant $-b_m^2 < 0$. Since $K_n(p_m, 0)$ is positive semi-definite according to Theorem 1.2 we have in both cases a contradiction. **QED**

Note that the quadratic polynomial t^2 is n-convex in any interval for all natural numbers n.

Corollary 1.5. *Let I be a finite interval, and let n be a natural number. There exists an n-convex function in I which is not $(n+1)$-convex in any subinterval of I.*

Note that the function in the corollary may be chosen as either an n-monotone increasing or an n-monotone decreasing polynomial of degree $2n$, and that the possible degrees of any polynomials in the gap are limited to $2n$ and $2n+1$.

Corollary 1.6. *Let I be an infinite interval different from the real line. For any natural number n there is an n-convex function in I which is not $(n+1)$-convex.*

Proof. We may without loss of generality assume that $I = [0, \infty)$. We may by Proposition 1.3 choose a polynomial f_n of degree $2n$ which in the interval $[0, 1)$ is n-monotone and n-concave. Possibly by adding a suitable constant we may assume that f_n is non-negative.

The transformation $t \to h(t) = t(1+t)^{-1}$ from $[0, \infty)$ to $[0, 1)$ is operator concave, therefore the function

$$g_n(t) = f_n\big(h(t)\big) = f_n\left(\frac{t}{1+t}\right) \qquad t \ge 0$$

is n-concave on $[0, \infty)$. But since the inverse transformation $t \to h^{-1}(t) = t(1-t)^{-1}$ from $[0, 1)$ to $[0, \infty)$ is operator monotone and $f_n = g_n \circ h^{-1}$ is not $(n+1)$-monotone, we derive that g_n is not $(n+1)$-monotone. But a non-negative $(n+1)$-concave function defined in the interval $[0, \infty)$ is necessarily $(n+1)$-monotone [4, Proposition 1.3]. We therefore conclude that g_n is not $(n+1)$-concave. **QED**

Note that the above proof does not exclude the possibility that g_n is $(n+1)$-concave in some subinterval of the half-line $[0, \infty)$.

2 Local property

We say that n-convexity is a local property if an arbitrary function f, defined in two overlapping open intervals I_1 and I_2 such that the restrictions of f to I_1 and I_2 are n-convex, necessarily is n-convex also in the union $I_1 \cup I_2$.

We conjecture that n-convexity, like n-monotonocity, is a local property, and we prove it for $n = 2$. The following representation of divided differences is due to Hermite [6].

Proposition 2.1. *Divided differences can be written in the following form*

$$\begin{aligned}
[x_0,x_1]_f &= \int_0^1 f'\Big((1-t_1)x_0+t_1x_1\Big)\,dt_1\\
[x_0,x_1,x_2]_f &= \int_0^1\int_0^{t_1} f''\Big((1-t_1)x_0+(t_1-t_2)x_1+t_2x_2\Big)\,dt_2\,dt_1\\
&\vdots\\
[x_0,x_1,\cdots,x_n]_f &= \int_0^1\int_0^{t_1}\cdots\int_0^{t_{n-1}} f^{(n)}\Big((1-t_1)x_0+(t_1-t_2)x_1+\cdots\\
&\qquad +(t_{n-1}-t_n)x_{n-1}+t_nx_n\Big)\,dt_n\cdots dt_2\,dt_1
\end{aligned}$$

where f is an n-times continuously differentiable function defined in an open interval I, and $x_0,x_1,\ldots,x_n$ are (not necessarily distinct) points in I.

2.1 An inequality for divided differences

Proposition 2.2. *Let I be an open interval and n a natural number. For a function $f\in C^n(I)$ we assume that the nth derivative $f^{(n)}$ is strictly positive. If in addition the function*

$$c(x)=\frac{1}{f^{(n)}(x)^{1/(n+1)}}\qquad t\in I$$

is convex, then the divided difference

$$[x_0,x_1,\ldots,x_n]_f\ge\prod_{i=0}^n\,[x_i,x_i,\ldots,x_i]_f^{1/(n+1)}$$

for arbitrary $x_0,x_1,\ldots,x_n\in I$, where the divided differences $[x_i,x_i,\ldots,x_i]_f$ are of order n. If on the other hand the (positive) function $c(x)$ is concave, then the inequality is reversed.

Proof. By using the expression for divided differences given in Proposition 2.1

and the convexity of the function c we obtain

$$[x_0, x_1, \cdots, x_n]_f = \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} f^{(n)}\Big((1-t_1)x_0 + (t_1-t_2)x_1 + \\ \cdots + (t_{n-1}-t_n)x_{n-1} + t_n x_n\Big)\, dt_n \cdots dt_2\, dt_1$$

$$= \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} c\Big((1-t_1)x_0 + (t_1-t_2)x_1 + \\ \cdots + (t_{n-1}-t_n)x_{n-1} + t_n x_n\Big)^{-(n+1)} dt_n \cdots dt_2\, dt_1$$

$$\geq \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} \Big((1-t_1)c(x_0) + (t_1-t_2)c(x_1) + \\ \cdots + (t_{n-1}-t_n)c(x_{n-1}) + t_n c(x_n)\Big)^{-(n+1)} dt_n \cdots dt_2\, dt_1.$$

Next considering the function

$$g(t) = \frac{1}{t} \qquad t > 0$$

with nth derivative

$$g^{(n)}(t) = (-1)^n \frac{n!}{t^{n+1}}$$

we may insert this in the above expression to obtain

$$[x_0, x_1, \ldots, x_n]_f \geq \frac{(-1)^n}{n!} \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} g^{(n)}\Big((1-t_1)c(x_0) + \\ (t_1-t_2)c(x_1) + \cdots + (t_{n-1}-t_n)c(x_{n-1}) + t_n c(x_n)\Big)\, dt_n \cdots dt_2\, dt_1$$

$$= \frac{(-1)^n}{n!}[c(x_0), c(x_1), \ldots, c(x_n)]_g$$

where we used Proposition 2.1 once more. Finally, since

$$[t_0, t_1, \ldots, t_n]_g = (-1)^n g(t_0) g(t_1) \cdots g(t_n) \qquad t_0, t_1, \ldots, t_n > 0$$

we obtain

$$\begin{aligned}
[x_0, x_1, \ldots, x_n]_f &\geq \frac{1}{n!}\, g(c(x_0)) g(c(x_1)) \cdots g(c(x_n)) \\
&= \frac{1}{n!\, c(x_0) c(x_1) \cdots c(x_n)} \\
&= \frac{1}{n!} f^{(n)}(x_0)^{1/(n+1)} f^{(n)}(x_1)^{1/(n+1)} \cdots f^{(n)}(x_n)^{1/(n+1)} \\
&= \prod_{i=0}^{n} \left(\frac{f^{(n)}(x_i)}{n!} \right)^{1/(n+1)}
\end{aligned}$$

and since $f^{(n)}(x) = n!\,[x, x, \ldots, x]_f$ for $x \in I$ the statement follows. If the function $c(x)$ is concave, the statement follows by making the appropriate alterations in the above proof. **QED**

We may use the above Proposition for the exponential function since the function

$$c(x) = \frac{1}{\exp^{(n)}(x)^{1/(n+1)}} = \exp(-x/(n+1))$$

indeed is convex. We therefore obtain

$$\begin{aligned}[x_0, x_1, \ldots, x_n]_{\exp} &\geq \prod_{i=0}^{n} [x_i, x_i, \ldots, x_i]_{\exp}^{1/(n+1)} \\ &= \frac{1}{n!} \exp\left(\frac{x_0 + x_1 + \cdots + x_n}{n+1}\right)\end{aligned}$$

for arbitrary real numbers $x_0, x_1, \ldots, x_n$. This consequence of Theorem 2.2 was proved in [8, Appendix 1] by another method.

2.2 Local property for 2-convex functions

Theorem 2.3. *Let I be an open interval, and take a function $f \in C^4(I)$ such that $f''(t) > 0$ for every $t \in I$. Then the following assertions are equivalent.*

(1) f is 2-convex.

(2) The matrix

$$\begin{pmatrix} \dfrac{f''(t)}{2} & \dfrac{f^{(3)}(t)}{6} \\ \dfrac{f^{(3)}(t)}{6} & \dfrac{f^{(4)}(t)}{24} \end{pmatrix}$$

is positive semi-definite for every $t \in I$.

(3) There is a positive concave function c on I such that $f''(t) = c(t)^{-3}$ for every $t \in I$.

(4) The inequality

$$[t_0, t_0, t_0]_f [t_1, t_1, t_1]_f - [t_0, t_1, t_1]_f [t_0, t_0, t_1]_f \geq 0$$

is valid for all $t_0, t_1 \in I$.

(5) The Kraus determinant

$$\begin{vmatrix} [t_0, t_0, t_0]_f & [t_0, t_0, t_1]_f \\ [t_0, t_0, t_1]_f & [t_0, t_1, t_1]_f \end{vmatrix} \geq 0$$

for all $t_0, t_1 \in I$.

Proof. (1) $\Rightarrow$ (2) is proved by Theorem 1.2.

(2) $\Rightarrow$ (3) : Put $c(t) = f''(t)^{-1/3}$ for $t \in I$. Then c is a positive function and $f''(t) = c(t)^{-3}$. By differentiation we obtain $f^{(3)}(t) = -3c(t)^{-4}c'(t)$ and

$$f^{(4)}(t) = 12c(t)^{-5}c'(t)^2 - 3c(t)^{-4}c''(t).$$

The determinant

$$\frac{f''(t)}{2}\frac{f^{(4)}(t)}{24} - \frac{f^{(3)}(t)}{36}^2$$

is non-negative by (2), thus inserting the derivatives we obtain

$$\begin{aligned} &\frac{c(t)^{-3}}{2} \cdot \frac{12c(t)^{-5}c'(t)^2 - 3c(t)^{-4}c''(t)}{24} - \frac{(-3c(t)^{-4}c'(t))^2}{36} \\ &= -\frac{1}{16}c(t)^{-7}c''(t) \geq 0, \end{aligned}$$

hence $c''(t) \leq 0$ for every $t \in I$ and c is concave.

(3) $\Rightarrow$ (4) : For $n = 2$ condition (3) becomes the assumption in Proposition 2.2, hence

$$[t_0, t_1, t_2]_f \leq [t_0, t_0, t_0]_f^{1/3}[t_1, t_1, t_1]_f^{1/3}[t_2, t_2, t_2]_f^{1/3}$$

for arbitrary $t_0, t_1, t_2 \in I$. Setting $t_2 = t_0$ we obtain

$$[t_0, t_0, t_1]_f \leq [t_0, t_0, t_0]_f^{2/3}[t_1, t_1, t_1]_f^{1/3}$$

and setting $t_2 = t_1$ we obtain

$$[t_0, t_1, t_1]_f \leq [t_0, t_0, t_0]_f^{1/3}[t_1, t_1, t_1]_f^{2/3},$$

hence the product

$$[t_0, t_1, t_1]_f[t_0, t_0, t_1]_f \leq [t_0, t_0, t_0]_f[t_1, t_1, t_1]_f$$

which is condition (4).

$(4) \Rightarrow (5)$: We introduce a function $F : I \to \mathbf{R}$ defined by setting $F(t_0) = 0$ and

$$\begin{aligned} F(t) = [t_0, t_0, t_0]_f \big((t - t_0) f'(t) - f(t) + f(t_0)\big) \\ - \frac{1}{(t - t_0)^2}((t_0 - t) f'(t_0) - f(t_0) + f(t))^2. \end{aligned}$$

for $t \neq t_0$. Since

$$\frac{1}{(t - t_0)^2}((t_0 - t) f'(t_0) - f(t_0) + f(t))^2 = (t - t_0)^2 [t_0, t_0, t]^2$$

for $t \neq t_0$ this defines F as a differentiable function, and since

$$[t_0, t, t]_f = \frac{1}{(t - t_0)^2}\Big((t - t_0) f'(t) - f(t) + f(t_0)\Big)$$

we obtain

$$[t_0, t_0, t_0]_f [t_0, t, t]_f - [t_0, t_0, t]_f^2 = \frac{1}{(t - t_0)^2} F(t) \tag{4}$$

for $t \neq t_0$. We next consider the derivative

$$\begin{aligned} F'(t) &= [t_0, t_0, t_0]_f \big(f'(t) + (t - t_0) f''(t) - f'(t)\big) \\ &\qquad + 2(t - t_0)^{-3}\big((t_0 - t) f'(t_0) - f(t_0) + f(t)\big)^2 \\ &\qquad - 2(t - t_0)^{-2}\big((t_0 - t) f'(t_0) - f(t_0) + f(t)\big)\big(-f'(t_0) + f'(t)\big) \\ &= 2(t - t_0)\Big([t_0, t_0, t_0]_f [t, t, t]_f - [t_0, t, t]_f [t_0, t_0, t]_f\Big). \end{aligned}$$

The assumption (4) entails that F has minimum in t_0 and therefore is non-negative. But this is equivalent to (5) by the identity (4).

$(5) \Rightarrow (1)$: This is the characterization by Kraus [7]. **QED**

Corollary 2.4. *2-convexity is a local property.*

Proof. By applying the regularization procedure described in the introduction we may assume that f is infinite many times differentiable. If f is an affine function there is nothing to prove. If f is not affine we may by Lemma 1.1 (2) assume that f'' is strictly positive. The statement is now a direct consequence of Theorem 2.3. **QED**

Corollary 2.5. *Let f be a twice continuously differentiable function defined in an open interval. If the determinant*

$$[t_0, t_0, t_0]_f [t_0, t_1, t_1]_f - [t_0, t_0, t_1]_f^2 = 0$$

for some $t_1 \neq t_0$ then also

$$[t_0, t_0, t_0]_f [t_0, t, t]_f - [t_0, t_0, t]_f^2 = 0$$

for any t between t_0 and t_1.

Proof. We first note that the implications (3) $\Rightarrow$ (4) and (4) $\Rightarrow$ (5) in the proof of Theorem 2.3 only require the function f to be twice continuously differentiable. The condition in the corollary entails that the function F defined in (4) for $t \neq t_0$ and with $F(t_0) = 0$ takes minimum both in t_0 and t_1 and consequently vanishes between the two points. **QED**

3 A fractional transformation

Let I be an open interval and take $t_0 \in I$. To each function $f \in C^2(I)$ such that $f'(t) > 0$ for every $t \in I$, Nayak [10] considered the following transformation

$$g_{t_0}(t) = -\frac{1}{f(t) - f(t_0)} + \frac{1}{f'(t_0)(t - t_0)}$$

which we write on the form

$$T(t_0, f)(t) = g_{t_0}(t) = \frac{[t_0, t_0, t]_f}{[t_0, t_0]_f [t_0, t]_f} \qquad t \in I. \tag{5}$$

The inverse transformation is given by

$$f(t) = f(t_0) - \cfrac{1}{T(t_0, f)(t) - \cfrac{1}{f'(t_0)(t - t_0)}} \tag{6}$$

for $t \neq t_0$ and $t \in I$. Nayak proved [10] the following result:

Theorem 3.1. *Let n be a natural number greater than or equal to two. The transform $T(t_0, f) \in P_n(I)$ for all $t_0 \in I$, if and only if $f \in P_{n+1}(I)$.*

Let I be an open interval and take $t_0 \in I$. To each function $f \in C^3(I)$ such that $f''(t) > 0$ for every $t \in I$, we consider the following transformation

$$S(t_0, f)(t) = \frac{[t_0, t_0, t_0, t]_f}{[t_0, t_0, t_0]_f [t_0, t_0, t]_f} \qquad t \in I \tag{7}$$

with inverse

$$f(t) = f(t_0) + f'(t_0)(t - t_0) - \frac{t - t_0}{S(t_0, f)(t) - \dfrac{1}{[t_0, t_0, t_0]_f (t - t_0)}} \tag{8}$$

for $t \neq t_0$ and $t \in I$. The two transformations are connected in the following way. Consider the function $d_{t_0}\colon I \to \mathbf{R}$ defined by setting $d_{t_0}(t) = [t_0, t]_f$. Since by a simple calculation

$$[t_0, t]_{d_{t_0}} = [t_0, t_0, t]_f \quad \text{and} \quad [t_0, t_0, t]_{d_{t_0}} = [t_0, t_0, t_0, t]_f$$

we obtain

$$S(t_0, f) = T(t_0, d_{t_0}). \tag{9}$$

But Nayak's result is not directly applicable since the function d_{t_0} depends on t_0.

Lemma 3.2. *Let $A = (a_{ij})_{i,j=0,1,\dots,k}$ be a $(k+1) \times (k+1)$ matrix and consider the $k \times k$ matrix $B = (b_{ij})_{i,j=1,\dots,k}$ defined by setting*

$$b_{ij} = a_{00} a_{ij} - a_{i0} a_{0j} \qquad i, j = 1, \dots, k.$$

Then the determinant $\det B = a_{00}^{k-1} \det A$.

Proof. We may express

$$b_{ij} = \det \begin{pmatrix} a_{00} & a_{0j} \\ a_{i0} & a_{ij} \end{pmatrix} \qquad i, j = 1, \dots, k$$

and observe that the result follows from Sylvester's determinant identity [1]. **QED**

Lemma 3.3. *The divided differences of the transform $S(t_0, f)$ in points $t_i, t_j \in I$ different from t_0 may be written on the form*

$$[t_i, t_j]_{S(t_0,f)} = \frac{[t_0, t_0]_{d_{t_0}} [t_i, t_j]_{d_{t_0}} - [t_i, t_0]_{d_{t_0}} [t_j, t_0]_{d_{t_0}}}{[t_0, t_0]_{d_{t_0}} (d_{t_0}(t_i) - d_{t_0}(t_0))(d_{t_0}(t_j) - d_{t_0}(t_0))}, \tag{10}$$

where as above the function $d_{t_0}\colon I \to \mathbf{R}$ is defined by setting $d_{t_0}(t) = [t_0, t]_f$.

Proof. By (9) and (5) we obtain

$$\begin{aligned} S(t_0, f)(t) &= \frac{[t_0, t_0, t]_{d_{t_0}}}{[t_0, t_0]_{d_{t_0}}[t_0, t]_{d_{t_0}}} = \frac{[t_0, t_0]_{d_{t_0}} - [t_0, t]_{d_{t_0}}}{[t_0, t_0]_{d_{t_0}}(d_{t_0}(t_0) - d_{t_0}(t))} \\ &= \frac{-1}{d_{t_0}(t) - d_{t_0}(t_0)} + \frac{1}{[t_0, t_0]_{d_{t_0}}(t - t_0)}. \end{aligned}$$

We calculate in distinct points t_i and t_j (different from t_0)

$$\begin{aligned} &(t_i - t_j)[t_i, t_j]_{S(t_0,f)} = S(t_0, f)(t_i) - S(t_0, f)(t_j) = \\ &\frac{-1}{d_{t_0}(t_i) - d_{t_0}(t_0)} + \frac{1}{[t_0, t_0]_{d_{t_0}}(t_i - t_0)} + \frac{1}{d_{t_0}(t_j) - d_{t_0}(t_0)} - \frac{1}{[t_0, t_0]_{d_{t_0}}(t_j - t_0)} \\ &= \frac{d_{t_0}(t_i) - d_{t_0}(t_j)}{(d_{t_0}(t_i) - d_{t_0}(t_0))(d_{t_0}(t_j) - d_{t_0}(t_0))} - \frac{t_i - t_j}{[t_0, t_0]_{d_{t_0}}(t_i - t_0)(t_j - t_0)} \end{aligned}$$

from which we obtain (10) and then realize that the identity holds in arbitrary points $t_i, t_j \in I$ different from t_0, which is the statement of the lemma. **QED**

Theorem 3.4. *Let $f \in C^3(I)$ where I is an open interval such that $f''(t) > 0$ for all $t \in I$, and let n be a natural number. Then the fractional transform $S(t_0, f)$ defined in (7) is in $P_n(I)$ for all $t_0 \in I$, if and only if $f \in K_{n+1}(I)$.*

Proof. Take a $t_0 \in I$ and a sequence $t_1, \ldots, t_n \in I$ of points different from t_0. For $k = 1, \ldots, n$ we calculate the determinant

$$\begin{aligned} &\det\left([t_i, t_j]_{S(t_0,f)}\right)_{i=1}^{k} \\ &= \sum_{\sigma \in S_k} \mathrm{Sign}(\sigma) \prod_{i=1}^{k} \left(\frac{[t_0, t_0]_{d_{t_0}}[t_i, t_{\sigma(i)}]_{d_{t_0}} - [t_i, t_0]_{d_{t_0}}[t_{\sigma(i)}, t_0]_{d_{t_0}}}{[t_0, t_0]_{d_{t_0}}(d_{t_0}(t_i) - d_{t_0}(t_0))(d_{t_0}(t_{\sigma(i)}) - d_{t_0}(t_0))} \right) \\ &= \frac{1}{[t_0, t_0]_{d_{t_0}}^k} \prod_{j=1}^{k} \frac{1}{(d_{t_0}(t_j) - d_{t_0}(t_0))^2} \sum_{\sigma \in S_k} \mathrm{Sign}(\sigma) \\ &\qquad \times \prod_{i=1}^{k} \left([t_0, t_0]_{d_{t_0}}[t_i, t_{\sigma(i)}]_{d_{t_0}} - [t_i, t_0]_{d_{t_0}}[t_{\sigma(i)}, t_0]_{d_{t_0}}\right) \\ &= \frac{\det B}{[t_0, t_0]_{d_{t_0}}^k} \prod_{j=1}^{k} \frac{1}{(d_{t_0}(t_j) - d_{t_0}(t_0))^2} \end{aligned}$$

where $B = (b_{ij})_{i,j=1}^k$ and

$$b_{ij} = [t_0, t_0]_{d_{t_0}}[t_i, t_j]_{d_{t_0}} - [t_i, t_0]_{d_{t_0}}[t_0, t_j]_{d_{t_0}}$$

for $i,j=1,\ldots,k$. If we set $A=(a_{ij})_{i,j=0}^k$ where $a_{ij}=[t_i,t_j]_{d_{t_0}}$ we may write

$$b_{ij}=a_{00}a_{ij}-a_{i0}a_{0j}\qquad i,j=1,\ldots,k.$$

By applying Lemma 3.2 we obtain

$$\det B=a_{00}^{k-1}\det A=[t_0,t_0]_{d_{t_0}}^{k-1}\det A$$

and by using the identity

$$\det A=\det\left([t_i,t_j]_{d_{t_0}}\right)_{i=0}^k=\det\left([t_0,t_i,t_j]_f\right)_{i,j=0}^k$$

we may write

$$\det\left([t_i,t_j]_{S(t_0,f)}\right)_{i,j=1}^k=\frac{\det\left([t_0,t_i,t_j]_f\right)_{i,j=0}^k}{[t_0,t_0]_{d_{t_0}}}\prod_{j=1}^k\frac{1}{(d_{t_0}(t_j)-d_{t_0}(t_0))^2}.$$

Suppose f is $(n+1)$-convex. We first obtain $\det\left([t_i,t_j]_{S(t_0,f)}\right)_{i,j=1}^k\geq 0$ for $k=1,\ldots,n$. By considering permutations of $t_1,\ldots,t_n$ we realize that the determinants of all the principal submatrices of each order k of the Pick matrix $\left([t_i,t_j]_{S(t_0,f)}\right)_{i,j=1}^n$ are non-negative, hence the Pick matrix itself is positive semi-definite. By continuity we obtain that the Pick matrix is positive semi-definite for arbitrary sequences $t_1,\ldots,t_n$ in I, hence $S(t_0,f)$ is n-monotone.

If on the other hand $S(t_0,f)$ is n-monotone, we realize that

$$\det\left([t_0,t_i,t_j]_f\right)_{i,j=0}^k\geq 0$$

for $k=1,\ldots,n$ and since for $k=0$ the entry $[t_0,t_0,t_0]_f>0$, we obtain that the leading determinants of the matrix $\left([t_0,t_i,t_j]_f\right)_{i,j=0}^k$ are non-negative. By considering permutations of $t_1,\ldots,t_n$ we realize that the determinants of all the principal submatrices of each order k of $\left([t_0,t_i,t_j]_f\right)_{i,j=0}^n$ are non-negative, hence the matrix is positive semi-definite. By continuity we finally realize that the Kraus matrix in equation (2) for f calculated in arbitrary $n+1$ points $t_0,t_1,\ldots,t_n\in I$ is positive semi-definite, hence f is $(n+1)$-convex. **QED**

References

[1] A.G. Akritas, E.K. Akritas, and G.I. Malaschonok. Various proofs of Sylvester's (determinant) identity. *Mathematics and Computers in Simulation*, 42:585–593, 1996.

[2] O. Dobsch. Matrixfunktionen beschränkter schwankung. *Math. Z.*, 43:353–388, 1937.

[3] W. Donoghue. *Monotone matrix functions and analytic continuation.* Springer, Berlin, Heidelberg, New York, 1974.

[4] F. Hansen. Operator monotone functions of several variables. *Math. Ineq. Appl.*, 6:1–17, 2003.

[5] F. Hansen, G. Ji, and J. Tomiyama. Gaps between classes of matrix monotone funstions. *Bull. London Math. Soc.*, 36:53–58, 2004.

[6] Ch. Hermite. Sur la formule d'interpolation de Lagrange. *Journal für die Reine und Angewandte Mathematik*, 84:70–79, 1878.

[7] F. Kraus. Über konvekse Matrixfunktionen. *Math. Z.*, 41:18–42, 1936.

[8] K.J. Le Couteur. Representation of the function $\mathrm{Tr}(\exp(a - \lambda b))$ as a Laplace transform with positive weigth and some matrix inequalities. *J. Phys. A: Math. Gen.*, 13:3147–3159, 1980.

[9] K. Löwner. Über monotone Matrixfunktionen. *Math. Z.*, 38:177–216, 1934.

[10] S. Nayak. Monotone matrix functions of succesive orders. *Proc. Amer. Math. Soc.*, 132:33–35, 2004.

[11] H. Osaka, S. Silvestrov, and J. Tomiyama. Monotone operator functions on C^*-algebras. *International Journal of Mathematics*, 16:181–196, 2005.

Frank Hansen: Department of Economics, University of Copenhagen, Studiestraede 6, DK-1455 Copenhagen K, Denmark.

Jun Tomiyama: Department of Mathematics and Physics. Japan Women's University. Mejirodai Bunkyo-ku, Tokyo, Japan.